

# Hölder-Type Global Error Bounds for Non-degenerate Polynomial Systems

Sĩ Tiệp ĐINH<sup>†</sup> · Huy Vui HÀ<sup>†</sup> · Tiến Sơn PHẠM<sup>‡</sup>

November 5, 2014

**Abstract** Let  $F := (f_1, \dots, f_p): \mathbb{R}^n \rightarrow \mathbb{R}^p$  be a polynomial map, and suppose that  $S := \{x \in \mathbb{R}^n : f_i(x) \leq 0, i = 1, \dots, p\} \neq \emptyset$ . Let  $d := \max_{i=1, \dots, p} \deg f_i$  and  $\mathcal{H}(d, n, p) := d(6d - 3)^{n+p-1}$ . Under the assumption that the map  $F: \mathbb{R}^n \rightarrow \mathbb{R}^p$  is convenient and non-degenerate at infinity, we show that there exists a constant  $c > 0$  such that the following so-called *Hölder-type global error bound result* holds

$$cd(x, S) \leq [f(x)]_+^{\frac{2}{\mathcal{H}(2d, n, p)}} + [f(x)]_+ \quad \text{for all } x \in \mathbb{R}^n,$$

where  $d(x, S)$  denotes the Euclidean distance between  $x$  and  $S$ ,  $f(x) := \max_{i=1, \dots, p} f_i(x)$ , and  $[f(x)]_+ := \max\{f(x), 0\}$ . The class of polynomial maps (with fixed Newton polyhedra), which are non-degenerate at infinity, is generic in the sense that it is an open and dense semi-algebraic set. Therefore, Hölder-type global error bounds hold for a large class of polynomial maps, which can be recognized relatively easily from their combinatoric data. This follows up the result on a Frank-Wolfe type theorem for non-degenerate polynomial programs in [16].

**Keywords** Error bounds; Newton polyhedron; non-degenerate polynomial maps; Palais-Smale condition.

**Mathematics Subject Classification (2010)** Primary 32B20; Secondary 14P, 49K40.

## 1 Introduction

Let  $f_1, \dots, f_p: \mathbb{R}^n \rightarrow \mathbb{R}$  be polynomial functions, and suppose that the set

$$S := \{x \in \mathbb{R}^n : f_1(x) \leq 0, \dots, f_p(x) \leq 0\}$$

---

Sĩ Tiệp ĐINH<sup>†</sup>

Institute of Mathematics, VAST, 18, Hoang Quoc Viet Road, Cau Giay District 10307, Hanoi, Vietnam

E-mail: dstiep@math.ac.vn

Huy Vui HÀ<sup>†</sup>

Institute of Mathematics, VAST, 18, Hoang Quoc Viet Road, Cau Giay District 10307, Hanoi, Vietnam

E-mail: hhvui@math.ac.vn

Tiến Sơn PHẠM<sup>‡</sup>

Department of Mathematics, University of Dalat, 1 Phu Dong Thien Vuong, Dalat, Vietnam

E-mail: sonpt@dlu.edu.vn

is nonempty. Let  $f(x) := \max_{i=1,\dots,p} f_i(x)$ . Then  $S = \{x \in \mathbb{R}^n : f(x) \leq 0\}$ . We are interested in the question of whether one can use the residual (constraint violation) at a point  $x \in \mathbb{R}^n$  to bound the distance from  $x$  to the set  $S$ . More precisely, we study if there exist some positive constants  $c, \alpha$ , and  $\beta$  such that

$$cd(x, S) \leq [f(x)]_+^\alpha + [f(x)]_+^\beta \quad \text{for all } x \in \mathbb{R}^n, \quad (1)$$

where  $d(x, S)$  denotes the Euclidean distance between  $x$  and the set  $S$  and  $[f(x)]_+ := \max\{f(x), 0\}$ . We say that a *Hölder-type global error bound* holds for the set  $S$  if the inequality (1) holds.

The study of error bounds has grown significantly and has found many important applications. In particular, it has been used sensitivity analysis for various problems of mathematical programming (for examples, the variational inequality, the linear and nonlinear complementarity problem, and the 0-1 integer feasibility problem). It has also been used as termination criteria for iterative decent algorithms. For a comprehensive, state of the art survey of the extensive theory and rich applications of error bounds, we refer the readers to the survey of Pang [49] and the references cited therein.

The first error bound result is due to Hoffman [25]. His result deals with the case where the polynomials  $f_1, \dots, f_p$  are affine and states that the inequality (1) holds with the exponents  $\alpha = \beta = 1$ . After the work of Hoffman, many people have devoted themselves to the study of global error bound; see, for example, [2], [31], [32], [35], [42], [47], [51].

In general, without the assumption of convexity, Hölder-type global error bounds are highly unlikely to hold. When the constrained set  $S$  defined by some affine functions and a *single quadratic* polynomial, Luo and Sturm [41] showed that the Hölder-type global error bound (1) holds with the exponents  $\alpha = \frac{1}{2}$  and  $\beta = 1$ . In particular, a Hölder-type global error bound (with some *unknown* fractional exponents  $\alpha$  and  $\beta$ ) was obtained very recently by Hà [24] for a nonlinear inequality defined by a *single* polynomial (i.e., in the case where  $p = 1$ ), which is convenient and non-degenerate at infinity (see [33] and Section 2 for precise definitions).

This paper will deal mainly with a class of polynomial maps, which are defined by combinatorial data and are called *non-degenerate at infinity*. This notion is an adaptation in the real setting of the notion of non-degeneracy in the complex setting given by [29], [33]. In both real and complex contexts, the class of polynomial maps (with fixed Newton polyhedra), which are non-degenerate at infinity, is generic in the sense that it is an open and dense set.

For any positive integers  $d, n$ , and  $p$ , let

$$\mathcal{H}(d, n, p) := d(6d - 3)^{n+p-1}.$$

With the definitions in the next section, the main contribution of this paper is the following Hölder-type global error bound with explicit exponents.

**Theorem 1** (Compare [24, Theorem C]) *Let  $F := (f_1, \dots, f_p): \mathbb{R}^n \rightarrow \mathbb{R}^p$ ,  $1 \leq p \leq n$ , be a polynomial map. Suppose that  $F$  is convenient and non-degenerate at infinity. Let  $f(x) := \max_{i=1,\dots,p} f_i(x)$  and  $S := \{x \in \mathbb{R}^n : f(x) \leq 0\} \neq \emptyset$ . Then there exists a constant  $c > 0$  such that*

$$cd(x, S) \leq [f(x)]_+^{\frac{2}{\mathcal{H}(2d, n, p)}} + [f(x)]_+ \quad \text{for all } x \in \mathbb{R}^n, \quad (2)$$

where  $d := \max_{i=1,\dots,p} \deg f_i$ .

Our result extends the result of [24], which studies the case  $p = 1$ . We also give estimations of the exponents  $\alpha, \beta$  in (1), which has not been done in [24]. In the case  $p = 1$ ,  $F = f_1$ , the existence of Hölder-type global error bounds follows easily from the existence of the following global Łojasiewicz-type inequality

$$cd(x, Z) \leq |F(x)|^\alpha + |F(x)|^\beta \quad \text{for all } x \in \mathbb{R}^n, \quad (3)$$

where  $Z = F^{-1}(0)$ . For a system of polynomials which is non-degenerate at infinity, the existence of (3) has been proved by [17]; however, the existence of Hölder-type global error bounds does not follow directly from (3). Our method is actually different from [24] at the crucial point that we use only the Curve Selection Lemma at infinity (see Lemma 1) as a tool. The reader may find other global versions of this inequality in papers [23], [26], [28].

It is worth notice that error bound results with *explicit* exponents are indeed important for both theory and applications since they can be used, e.g., to establish explicit *convergence rates* of the proximal point algorithm as demonstrated in [9], [37], [38].

Note that we do not impose the condition of convexity on the polynomials  $f_i$ , and their degrees can be arbitrary. Further, by genericity of the condition of non-degeneracy at infinity, the Hölder-type global error bounds hold for almost polynomial maps.

The above Hölder-type global error bound result, together with the Frank-Wolfe type theorem in [16], suggests that the class of polynomial maps, which are non-degenerate at infinity, may offer an appropriate domain on which the machinery of polynomial optimization works with full efficiency.

Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  be a continuous semi-algebraic function. Assume that  $S := \{x \in \mathbb{R}^n : f(x) \leq 0\} \neq \emptyset$ . In order to obtain the main theorem, we have established some intermediate results which are of independent interest. Our principal tool is Curve Selection Lemma at infinity and the proof is closed to [16], [17]. The sketch of the proof of the main result is as follows.

- First of all, we provide a necessary and sufficient condition for the existence of a Hölder-type global error bound for the set  $S$  (Theorem 2).
- Secondly, we show that if  $f$  satisfies the *Palais-Smale condition* at each non-negative value, then  $f$  satisfies the above sufficient condition (Theorem 3).
- Thirdly, we show that if  $f$  has a *good asymptotic behavior* at infinity, then a Hölder-type global error bound (with the exponent  $\beta = 1$ ) holds for the set  $S$  (Theorem 4).
- Finally, we prove that if the map  $(f_1, \dots, f_p): \mathbb{R}^n \rightarrow \mathbb{R}^p$  is convenient and non-degenerate at infinity, then  $f(x) := \max_{i=1, \dots, p} f_i(x)$  has a good asymptotic behavior at infinity (Lemma 7). Specifically, we determine explicitly the exponents  $\alpha$  and  $\beta$  in the Hölder-type global error bound (1) for such polynomial systems.

The paper is structured as follows. Section 2 presents some backgrounds in the field. A criterion for the existence of a Hölder-type global error bound is given in Section 3. A relation between the Palais-Smale condition and the existence of Hölder-type global error bounds is given in Section 4. In Section 5, we consider goodness at infinity. The Hölder-type global error bound result (Theorem 1) for convenient and non-degenerate polynomial maps will be proven in Section 6. In Section 7, we give some illustrated examples for Theorem 1. Section 8 presents some applications.

## 2 Preliminaries

### 2.1 Semi-algebraic geometry

In this subsection, we recall some notions and results of semi-algebraic geometry, which can be found in [5], [6], [7], [18].

**Definition 1** (i) A subset of  $\mathbb{R}^n$  is called *semi-algebraic* if it is a finite union of sets of the form

$$\{x \in \mathbb{R}^n : f_i(x) = 0, i = 1, \dots, k; f_i(x) > 0, i = k + 1, \dots, p\}$$

where all  $f_i$  are polynomials.

(ii) Let  $A \subset \mathbb{R}^n$  and  $B \subset \mathbb{R}^p$  be semi-algebraic sets. A map  $F: A \rightarrow B$  is said to be *semi-algebraic* if its graph

$$\{(x, y) \in A \times B : y = F(x)\}$$

is a semi-algebraic subset in  $\mathbb{R}^n \times \mathbb{R}^p$ .

We list below some basic properties of semi-algebraic sets and functions.

- (i) The class of semi-algebraic sets is closed under Boolean operators, taking Cartesian product, closure and interior.
- (ii) A composition of semi-algebraic maps is a semi-algebraic map; the image and preimage of a semi-algebraic set under a semi-algebraic map are semi-algebraic sets;
- (iii) If  $S$  is a semi-algebraic set, then the distance function

$$d(\cdot, S): \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto d(x, S) := \inf\{\|x - a\| : a \in S\},$$

is also semi-algebraic.

We give a version of the Curve Selection Lemma which will be used in the paper. For more details, see [44], [46] and see [16] for a complete proof.

**Lemma 1 (Curve Selection Lemma at infinity)** *Let  $A \subset \mathbb{R}^n$  be a semi-algebraic set, and let  $F := (f_1, \dots, f_p): \mathbb{R}^n \rightarrow \mathbb{R}^p$  be a semi-algebraic map. Assume that there exists a sequence  $x^k \in A$  such that  $\lim_{k \rightarrow \infty} \|x^k\| = \infty$  and  $\lim_{k \rightarrow \infty} F(x^k) = y \in (\overline{\mathbb{R}})^p$ , where  $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ . Then there exists a smooth semi-algebraic curve  $\varphi: (0, \epsilon) \rightarrow \mathbb{R}^n$  such that  $\varphi(t) \in A$  for all  $t \in (0, \epsilon)$ ,  $\lim_{t \rightarrow 0} \|\varphi(t)\| = \infty$ , and  $\lim_{t \rightarrow 0} F(\varphi(t)) = y$ .*

The following result is useful in the next section (see, e.g., [18], [43]).

**Lemma 2 (Growth Dichotomy Lemma)** *Let  $f: (0, \epsilon) \rightarrow \mathbb{R}$  be a semi-algebraic function with  $f(t) \neq 0$  for all  $t \in (0, \epsilon)$ . Then there exist constants  $c \neq 0$  and  $q \in \mathbb{Q}$  such that  $f(t) = ct^q + o(t^q)$  as  $t \rightarrow 0^+$ .*

### 2.2 Newton polyhedra

In many problems, the combinatorial informations of polynomial maps are important and can be found in their Newton polyhedra. In this subsection, we recall the definition of Newton polyhedra following Kouchnirenko and Khovanskii (see [33], [29]).

Let us begin with some notations which will be used throughout this work. We consider a fixed coordinate system  $x_1, \dots, x_n \in \mathbb{R}^n$ . Let  $J \subset \{1, \dots, n\}$ , then we define

$$\mathbb{R}^J := \{x \in \mathbb{R}^n : x_j = 0, \text{ for all } j \notin J\}.$$

We denote by  $\mathbb{R}_{\geq 0}$  the set of non-negative real numbers. We also set  $\mathbb{Z}_{\geq 0} := \mathbb{R}_{\geq 0} \cap \mathbb{Z}$ . If  $\kappa = (\kappa_1, \dots, \kappa_n) \in \mathbb{Z}_{\geq 0}^n$ , we denote by  $x^\kappa$  the monomial  $x_1^{\kappa_1} \cdots x_n^{\kappa_n}$  and by  $|\kappa|$  the sum  $\kappa_1 + \cdots + \kappa_n$ .

**Definition 2** A subset  $\Gamma \subset \mathbb{R}_{\geq 0}^n$  is said to be a *Newton polyhedron at infinity*, if there exists some finite subset  $A \subset \mathbb{Z}_{\geq 0}^n$  such that  $\Gamma$  is equal to the convex hull in  $\mathbb{R}^n$  of  $A \cup \{0\}$ . Hence we say that  $\Gamma$  is the Newton polyhedron at infinity determined by  $A$  and we write  $\Gamma = \Gamma(A)$ . We say that a Newton polyhedron at infinity  $\Gamma \subset \mathbb{R}_{\geq 0}^n$  is *convenient* if it intersects each coordinate axis in a point different from the origin, that is, if for any  $i \in \{1, \dots, n\}$  there exists some integer  $m_i > 0$  such that  $m_i e_i \in \Gamma$ , where  $\{e_1, \dots, e_n\}$  denotes the canonical basis in  $\mathbb{R}^n$ .

Given a Newton polyhedron at infinity  $\Gamma \subset \mathbb{R}_{\geq 0}^n$  and a vector  $q \in \mathbb{R}^n$ , we define

$$\begin{aligned} d(q, \Gamma) &:= \min\{\langle q, \kappa \rangle : \kappa \in \Gamma\}, \\ \Delta(q, \Gamma) &:= \{\kappa \in \Gamma : \langle q, \kappa \rangle = d(q, \Gamma)\}. \end{aligned}$$

We say that a subset  $\Delta$  of  $\Gamma$  is a *face* of  $\Gamma$  if there exists a vector  $q \in \mathbb{R}^n$  such that  $\Delta = \Delta(q, \Gamma)$ . The dimension of a face  $\Delta$  is defined as the minimum of the dimensions of the affine subspaces containing  $\Delta$ . The faces of  $\Gamma$  of dimension 0 are called the *vertices* of  $\Gamma$ . We denote by  $\Gamma_\infty$  the set of the faces of  $\Gamma$  which do not contain the origin 0 in  $\mathbb{R}^n$ .

Let  $\Gamma_1, \dots, \Gamma_p$  be a collection of  $p$  Newton polyhedra at infinity in  $\mathbb{R}_{\geq 0}^n$ , for some  $p \geq 1$ . The *Minkowski sum* of  $\Gamma_1, \dots, \Gamma_p$  is defined as the set

$$\Gamma_1 + \cdots + \Gamma_p = \{\kappa^1 + \cdots + \kappa^p : \kappa^i \in \Gamma_i, \text{ for all } i = 1, \dots, p\}.$$

By definition,  $\Gamma_1 + \cdots + \Gamma_p$  is again a Newton polyhedron at infinity. Moreover, by applying the definitions given above, it is easy to check that

$$\begin{aligned} d(q, \Gamma_1 + \cdots + \Gamma_p) &= d(q, \Gamma_1) + \cdots + d(q, \Gamma_p), \\ \Delta(q, \Gamma_1 + \cdots + \Gamma_p) &= \Delta(q, \Gamma_1) + \cdots + \Delta(q, \Gamma_p), \end{aligned}$$

for all  $q \in \mathbb{R}^n$ . As an application of these relations, we obtain the following lemma whose proof can be found in [16].

**Lemma 3** (i) Assume that  $\Gamma$  is a convenient Newton polyhedron at infinity. Let  $\Delta$  be a face of  $\Gamma$  and let  $q = (q_1, \dots, q_n) \in \mathbb{R}^n$  such that  $\Delta = \Delta(q, \Gamma)$ . Then the following conditions are equivalent:

- (i1)  $\Delta \in \Gamma_\infty$ ;
- (i2)  $d(q, \Gamma) < 0$ ;
- (i3)  $\min_{j=1, \dots, n} q_j < 0$ .

(ii) Assume that  $\Gamma_1, \dots, \Gamma_p$  are some Newton polyhedra at infinity. Let  $\Delta$  be a face of the Minkowski sum  $\Gamma := \Gamma_1 + \cdots + \Gamma_p$ . Then the following statements hold:

- (ii1) There exists a unique collection of faces  $\Delta_1, \dots, \Delta_p$  of  $\Gamma_1, \dots, \Gamma_p$ , respectively, such that

$$\Delta = \Delta_1 + \cdots + \Delta_p.$$

(ii2) If  $\Gamma_1, \dots, \Gamma_p$  are convenient, then  $\Gamma_\infty \subset \Gamma_{1,\infty} + \dots + \Gamma_{p,\infty}$ .

Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  be a polynomial function. Suppose that  $f$  is written as  $f = \sum_{\kappa} a_{\kappa} x^{\kappa}$ . Then the support of  $f$ , denoted by  $\text{supp}(f)$ , is defined as the set of those  $\kappa \in \mathbb{Z}_{\geq 0}^n$  such that  $a_{\kappa} \neq 0$ . We denote the set  $\Gamma(\text{supp}(f))$  by  $\Gamma(f)$ . This set will be called the *Newton polyhedron at infinity* of  $f$ . The polynomial  $f$  is said to be convenient when  $\Gamma(f)$  is convenient. If  $f \equiv 0$ , then we set  $\Gamma(f) = \emptyset$ . Note that, if  $f$  is convenient, then for each nonempty subset  $J$  of  $\{1, \dots, n\}$ , we have  $\Gamma(f) \cap \mathbb{R}^J = \Gamma(f|_{\mathbb{R}^J})$ . The *Newton boundary at infinity* of  $f$ , denoted by  $\Gamma_\infty(f)$ , is defined as the set of the faces of  $\Gamma(f)$  which do not contain the origin  $0$  in  $\mathbb{R}^n$ .

Let us fix a face  $\Delta$  of  $\Gamma_\infty(f)$ . We define the *principal part of  $f$  at infinity with respect to  $\Delta$* , denoted by  $f_\Delta$ , as the sum of those terms  $a_{\kappa} x^{\kappa}$  such that  $\kappa \in \Delta$ .

*Remark 1* By definition, for each face  $\Delta$  of  $\Gamma_\infty$  there exists a vector  $q = (q_1, \dots, q_n) \in \mathbb{R}^n$  with  $\min_{j=1, \dots, n} q_j < 0$  such that  $\Delta = \Delta(q, \Gamma)$ .

### 2.3 Non-degeneracy at infinity

In [29] (see also [33]), Khovanskii introduced a condition of non-degeneracy of complex analytic maps  $F: (\mathbb{C}^n, 0) \rightarrow (\mathbb{C}^p, 0)$  in terms of the Newton polyhedra of the component functions of  $F$ . The class of non-degenerate maps is sufficiently large and plays an important role in Singularity Theory and Algebraic Geometry (see, for instance, [1], [21], [48]). We will apply this condition for real polynomial maps. First we need to introduce some notations.

Let  $F := (f_1, \dots, f_p): \mathbb{R}^n \rightarrow \mathbb{R}^p, 1 \leq p \leq n$ , be a polynomial map. Let  $\Gamma(F)$  denote the Minkowski sum  $\Gamma(f_1) + \dots + \Gamma(f_p)$ , and we denote by  $\Gamma_\infty(F)$  the set of faces of  $\Gamma(F)$  which do not contain the origin  $0$  in  $\mathbb{R}^n$ . Let  $\Delta$  be a face of the  $\Gamma(F)$ . According to Lemma 3, let us consider the decomposition  $\Delta = \Delta_1 + \dots + \Delta_p$ , where  $\Delta_i$  is a face of  $\Gamma(f_i)$ , for all  $i = 1, \dots, p$ . We denote by  $F_\Delta$  the polynomial map  $(f_{1,\Delta_1}, \dots, f_{p,\Delta_p}): \mathbb{R}^n \rightarrow \mathbb{R}^p$ .

**Definition 3** The polynomial map  $F = (f_1, \dots, f_p)$  is called *convenient* if all  $f_i$  are convenient, for  $i = 1, \dots, p$ .

We say that  $F$  is *non-degenerate at infinity* if and only if for any face  $\Delta$  of  $\Gamma_\infty(F)$  and for all  $x \in (\mathbb{R} \setminus \{0\})^n$ , we have  $\text{rank} M_\Delta = p$  where

$$M_\Delta := \begin{pmatrix} x_1 \frac{\partial f_{1,\Delta_1}}{\partial x_1}(x) & \dots & x_n \frac{\partial f_{1,\Delta_1}}{\partial x_n}(x) & f_{1,\Delta_1}(x) & \dots & 0 \\ \vdots & \dots & \vdots & \dots & \ddots & \vdots \\ x_1 \frac{\partial f_{p,\Delta_p}}{\partial x_1}(x) & \dots & x_n \frac{\partial f_{p,\Delta_p}}{\partial x_n}(x) & 0 & \dots & f_{p,\Delta_p}(x) \end{pmatrix}.$$

*Remark 2* Compared to  $F = (f_1, \dots, f_p)$ , the polynomial map  $F_\Delta = (f_{1,\Delta_1}, \dots, f_{p,\Delta_p})$  has the following two remarkable properties:

- The sparsity, which means that the number of monomials in  $f_{i,\Delta_i}$  is much less than that in  $f_i$ . By the Khovanskii's theory of fewnomials, see [30], working with  $F_\Delta$  is easier than with  $F$ .
- $F_\Delta$  is quasi-homogeneous, i.e., there exists a vector  $q \in \mathbb{R}^n$ , with  $\min_j q_j < 0$ , such that for each  $i = 1, \dots, p$ , we have  $\Delta_i = \Delta_i(q, \Gamma(f_i))$  and

$$f_{i,\Delta_i}(t^{q_1} x_1, \dots, t^{q_n} x_n) = t^{d_i} f_{i,\Delta_i}(x_1, \dots, x_n),$$

for all  $t \in \mathbb{R}$ , and for all  $(x_1, \dots, x_n) \in \mathbb{R}^n$ , where  $d_i := d(q, \Gamma(f_i))$ .

These facts, in many contexts, allows us to check easily the non-degenerate condition.

### 3 The existence of a Hölder-type global error bound

In this section we give a necessary and sufficient condition for the existence of a Hölder-type global error bound for a semi-algebraic function. This result extends Theorem A of [24] from polynomial functions to semi-algebraic functions.

Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  be a continuous semi-algebraic function. Assume that  $S := \{x \in \mathbb{R}^n : f(x) \leq 0\} \neq \emptyset$ . Let  $[f(x)]_+ := \max\{f(x), 0\}$ .

**Theorem 2** *With the notations above, the following two statements are equivalent.*

- (i) *For any sequence  $x^k \in \mathbb{R}^n \setminus S, x^k \rightarrow \infty$ , we have*
  - (i1) *if  $f(x^k) \rightarrow 0$  then  $d(x^k, S) \rightarrow 0$ ;*
  - (i2) *if  $d(x^k, S) \rightarrow \infty$  then  $f(x^k) \rightarrow \infty$ .*
- (ii) *There exist some constants  $c > 0, \alpha > 0$ , and  $\beta > 0$  such that*

$$cd(x, S) \leq [f(x)]_+^\alpha + [f(x)]_+^\beta \quad \text{for all } x \in \mathbb{R}^n.$$

The proof is essentially the same as the proof of [24, Theorem A] (see also [15, Proposition 3.10]), in fact, the theorem follows from the next two lemmas that we leave the reader verifying the details.

**Lemma 4 (Hölder-type error bound “near to  $S$ ”)** *The following two statements are equivalent.*

- (i) *For any sequence  $x^k \in \mathbb{R}^n \setminus S$ , with  $x^k \rightarrow \infty$ , it holds that*

$$f(x^k) \rightarrow 0 \implies d(x^k, S) \rightarrow 0;$$

- (ii) *There exist some constants  $c > 0, \delta > 0$ , and  $\alpha > 0$  such that*

$$cd(x, S) \leq [f(x)]_+^\alpha \quad \text{for all } x \in f^{-1}((-\infty, \delta]).$$

**Lemma 5 (Hölder-type error bound “far from  $S$ ”)** *Suppose that for any sequence  $x^k \in \mathbb{R}^n \setminus S$ , with  $x^k \rightarrow \infty$ , it holds that*

$$d(x^k, S) \rightarrow \infty \implies f(x^k) \rightarrow \infty;$$

*Then there exist some constants  $c > 0, r > 0$ , and  $\beta > 0$  such that*

$$cd(x, S) \leq [f(x)]_+^\beta \quad \text{for all } x \in f^{-1}([r, +\infty)).$$

### 4 The Palais-Smale condition and Hölder-type global error bounds

The relation between the Palais-Smale condition and the existence of error bounds is well-known, see, for example, [3], [4], [14], [24], [27], [34], [50]. In this part, we describe this relation in a more convenient form for our present purposes.

First of all, we recall the notion of subdifferential—that is, an appropriate multivalued operator playing the role of the usual gradient map—which is crucial for our considerations. For nonsmooth analysis we refer to the comprehensive texts [12], [13], [45], [52].

**Definition 4** (i) The Fréchet subdifferential  $\hat{\partial}f(x)$  of a continuous function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  at  $x \in \mathbb{R}^n$  is given by

$$\hat{\partial}f(x) := \left\{ v \in \mathbb{R}^n : \liminf_{\|h\| \rightarrow 0, h \neq 0} \frac{f(x+h) - f(x) - \langle v, h \rangle}{\|h\|} \geq 0 \right\}.$$

- (ii) The *limiting subdifferential* at  $x \in \mathbb{R}^n$ , denoted by  $\partial f(x)$ , is the set of all cluster points of sequences  $\{v^k\}_{k \geq 1}$  such that  $v^k \in \hat{\partial} f(x^k)$  and  $(x^k, f(x^k)) \rightarrow (x, f(x))$  as  $k \rightarrow \infty$ .

*Remark 3* It is a well-known result of variational analysis that  $\hat{\partial} f(x)$  (and a fortiori  $\partial f(x)$ ) is not empty in a dense subset of the domain of  $f$  (see [52], for example).

**Definition 5** Using the limiting subdifferential  $\partial f$ , we define the *nonsmooth slope* of  $f$  by

$$\mathbf{m}_f(x) := \inf\{\|v\| : v \in \partial f(x)\}.$$

By definition,  $\mathbf{m}_f(x) = +\infty$  whenever  $\partial f(x) = \emptyset$ .

*Remark 4* (i) If the function  $f$  is of class  $C^1$ , the above notion coincides with the usual concept of gradient; that is,  $\partial f(x) = \hat{\partial} f(x) = \{\nabla f(x)\}$ , and hence  $\mathbf{m}_f(x) = \|\nabla f(x)\|$ .

(ii) By Tarski-Seidenberg Theorem (see [5], [7]), it is not hard to show that if the function  $f$  is semi-algebraic then so is  $\mathbf{m}_f$ .

**Definition 6** Given a continuous function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  and a real number  $t$ , we say that  $f$  satisfies the *Palais-Smale condition at the level  $t$* , if every sequence  $\{x^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$  such that  $f(x^k) \rightarrow t$  and  $\mathbf{m}_f(x^k) \rightarrow 0$  as  $k \rightarrow \infty$  possesses a convergence subsequence.

The following result extends [24, Theorem B] from polynomial functions to semi-algebraic functions.

**Theorem 3** Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  be a continuous semi-algebraic function. Assume that  $S := \{x \in \mathbb{R}^n : f(x) \leq 0\} \neq \emptyset$ . If  $f$  satisfies the Palais-Smale condition at each level  $t \geq 0$ , then there exist some constants  $c > 0$ ,  $\alpha > 0$ , and  $\beta > 0$  such that

$$cd(x, S) \leq [f(x)]_+^\alpha + [f(x)]_+^\beta \quad \text{for all } x \in \mathbb{R}^n.$$

*Proof* The proof is similar to that of [24, Theorem B]. However, instead of using the Ekeland Variational Principle [19], we use a version of the variational principle of Borwein and Preiss (see [10], [13, Theorem 4.2]).

It is sufficient to show that the condition (i) in Theorem 2 holds. We proceed by the method of contradiction.

We first assume that there exist a number  $\delta > 0$  and a sequence  $x^k \in \mathbb{R}^n \setminus S$ , with  $x^k \rightarrow \infty$ , such that

$$f(x^k) \rightarrow 0 \quad \text{and} \quad d(x^k, S) \geq \delta.$$

Let us consider the continuous semi-algebraic function

$$f_+: \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \max\{f(x), 0\}.$$

Clearly,  $\inf_{x \in \mathbb{R}^n} f_+(x) = 0$ . Applying the Minimization Principle [13, Theorem 4.2] to the function  $f_+$  with data  $\epsilon := f_+(x^k) = f(x^k) > 0$  and  $\lambda := \frac{\delta}{4} > 0$ , we find points  $y^k$  and  $z^k$  in  $\mathbb{R}^n$  such that

$$\|z^k - x^k\| < \lambda, \quad \|y^k - z^k\| < \lambda, \quad f_+(y^k) \leq f_+(x^k),$$

and such that the function

$$x \mapsto f_+(x) + \frac{\epsilon}{\lambda^2} \|x - z^k\|^2$$



is minimized over  $\mathbb{R}^n$  at  $y^k$ . We deduce from the above inequalities that

$$\|y^k - x^k\| \leq \|z^k - x^k\| + \|y^k - z^k\| < 2\lambda = \frac{\delta}{2},$$

which yields that  $\lim_{k \rightarrow \infty} \|y^k\| = \infty$  and

$$d(y^k, S) \geq d(x^k, S) - d(x^k, y^k) > d(x^k, S) - \frac{\delta}{2} \geq \frac{\delta}{2}.$$

Hence,

$$B(y^k, \frac{\delta}{2}) := \left\{ x \in \mathbb{R}^n : \|x - y^k\| < \frac{\delta}{2} \right\} \subset \mathbb{R}^n \setminus S.$$

In particular, we have  $f_+(x) = f(x)$  for all  $x \in B(y^k, \frac{\delta}{2})$ . Consequently, the function

$$x \mapsto f(x) + \frac{\epsilon}{\lambda^2} \|x - z^k\|^2$$

attains its minimum on the open ball  $B(y^k, \frac{\delta}{2})$  at  $y^k$ . Then, by the Fermat's rule generalized [52, Theorem 10.1], we get

$$-2\frac{\epsilon}{\lambda^2}(y^k - z^k) \in \partial f(y^k).$$

Therefore

$$\mathbf{m}_f(y^k) \leq 2\frac{\epsilon}{\lambda^2} \|y^k - z^k\| \leq 2\frac{\epsilon}{\lambda} \leq \frac{8f(x^k)}{\delta}.$$

By letting  $k$  tend to infinity, we obtain

$$\lim_{k \rightarrow \infty} \|y^k\| = \infty, \quad \lim_{k \rightarrow \infty} f(y^k) = 0, \quad \text{and} \quad \lim_{k \rightarrow \infty} \mathbf{m}_f(y^k) = 0.$$

So,  $f$  does not satisfy the Palais-Smale condition at the value  $t = 0$ , and a contradiction follows.

We next suppose that there exist a number  $M > 0$  and a sequence  $x^k \in \mathbb{R}^n \setminus S$ , with  $x^k \rightarrow \infty$ , such that

$$d(x^k, S) \rightarrow \infty \quad \text{and} \quad 0 < f(x^k) \leq M.$$

Again, we see that  $\inf_{x \in \mathbb{R}^n} f_+(x) = 0$ . We now apply the Minimization Principle [13, Theorem 4.2] to the function  $f_+$  with data  $\epsilon := f_+(x^k) = f(x^k) > 0$  and  $\lambda := \frac{d(x^k, S)}{4} > 0$ ; there exist points  $y^k$  and  $z^k$  in  $\mathbb{R}^n$  with

$$\|z^k - x^k\| < \lambda, \quad \|y^k - z^k\| < \lambda, \quad f_+(y^k) \leq f_+(x^k),$$

and having the property that the function

$$x \mapsto f_+(x) + \frac{\epsilon}{\lambda^2} \|x - z^k\|^2$$

has a unique minimum at  $y^k$ . We deduce from the above inequalities that

$$\begin{aligned} d(y^k, S) &\geq d(x^k, S) - d(x^k, y^k) \\ &\geq d(x^k, S) - 2\lambda = \frac{d(x^k, S)}{2}, \end{aligned}$$

which yields  $\lim_{k \rightarrow \infty} d(y^k, S) = \infty$ . In particular, we get  $y^k \in \mathbb{R}^n \setminus S$  and  $y^k \rightarrow \infty$ .

By an argument as above, we can easily deduce again that

$$\mathbf{m}_f(y^k) \leq 2\frac{\epsilon}{\lambda^2} \|y^k - z^k\| \leq 2\frac{\epsilon}{\lambda} \leq \frac{8M}{d(x^k, S)}.$$

Hence,

$$\lim_{k \rightarrow \infty} \mathbf{m}_f(y^k) = 0.$$

Note that  $0 < f(y^k) \leq f(x^k) \leq M$  for all  $k \geq 1$ . Hence, by passing to subsequences if necessary, we may assume that there exists the limit  $t := \lim_{k \rightarrow \infty} f(y^k)$ . Therefore  $f$  does not satisfy the Palais-Smale condition at  $t$ , which is a contradiction. The proof of Theorem 3 is complete.

## 5 Goodness at infinity and Hölder-type global error bounds

The aim of this section is to establish a Hölder-type global error bound (with the exponent  $\beta = 1$ ) for continuous semi-algebraic functions which have good asymptotic behavior at infinity. Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  be a continuous function. For  $x \in \mathbb{R}^n$ , set

$$f_+(x) := \max\{f(x), 0\}.$$

Then  $f_+$  is also a continuous function. Let us begin with the following definition.

**Definition 7** A continuous function  $f$  is said to be *good at infinity* if there exist some constants  $c > 0$  and  $R > 0$  such that

$$\mathbf{m}_f(x) \geq c \quad \text{for all } x \in f^{-1}((0, +\infty)) \text{ and } \|x\| \geq R.$$

The main result of this section is as follows:

**Theorem 4** Let  $f$  be a continuous semi-algebraic function which is good at infinity. Assume that  $S := \{x \in \mathbb{R}^n : f(x) \leq 0\} \neq \emptyset$ . Then there exist some constants  $c > 0$  and  $\alpha > 0$  such that

$$cd(x, S) \leq [f(x)]_+^\alpha + [f(x)]_+ \quad \text{for all } x \in \mathbb{R}^n.$$

*Proof* Let us consider the continuous semi-algebraic function

$$f_+: \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \max\{f(x), 0\}.$$

By definition, if  $f(x) > 0$  then  $f_+(x) = f(x)$ ,  $\partial f_+(x) = \partial f(x)$  and  $\mathbf{m}_{f_+}(x) = \mathbf{m}_f(x)$ . Since  $f$  is good at infinity, there exist some constants  $c_1 > 0$  and  $R > 0$  such that

$$\mathbf{m}_{f_+}(x) = \mathbf{m}_f(x) \geq c_1 \quad \text{for all } x \in f^{-1}((0, +\infty)) = f_+^{-1}((0, +\infty)) \text{ and } \|x\| \geq R. \quad (4)$$

Thanks to the classical Lojasiewicz inequality [26], [39], there are constants  $c_2 > 0$  and  $\alpha > 0$  such that

$$c_2 d(x, S) \leq f_+(x)^\alpha \quad \text{for all } \|x\| \leq R. \quad (5)$$

Let  $x \in \mathbb{R}^n$  be such that  $x \in f_+^{-1}((0, +\infty))$  and  $\|x\| > R$ . By [8, Corollary 4.1], there exists a maximal absolutely continuous curve  $u: [0, \infty) \rightarrow \mathbb{R}^n$  of the dynamical system

$$0 \in \dot{u}(s) + \partial[f_+(u(s))]$$

satisfying  $u(0) = x$ . In addition, the function  $s \mapsto (f_+ \circ u)(s)$  is absolutely continuous and strictly decreasing on  $[0, +\infty)$ . By [8, Corollary 4.2], we have for almost all  $s \in [0, +\infty)$ ,

$$\|\dot{u}(s)\| = \mathbf{m}_{f_+}(u(s)) \quad \text{and} \quad \frac{d}{ds}(f_+ \circ u)(s) = -[\mathbf{m}_{f_+}(u(s))]^2. \quad (6)$$

We have the following remark. Suppose that  $f_+(u(s)) > 0$  and  $\|u(s)\| \geq R$  for all  $s \in [t_1, t_2]$ , for some  $0 \leq t_1 < t_2$ . It follows from the relations (4) and (6) that

$$\begin{aligned} f_+(u(t_1)) - f_+(u(t_2)) &= - \int_{t_1}^{t_2} \frac{d}{ds}(f_+ \circ u)(s) ds = \int_{t_1}^{t_2} [\mathbf{m}_{f_+}(u(s))]^2 ds \\ &\geq \int_{t_1}^{t_2} c_1 \mathbf{m}_{f_+}(u(s)) ds = \int_{t_1}^{t_2} c_1 \|\dot{u}(s)\| ds, \end{aligned}$$

which yields

$$f_+(u(t_1)) - f_+(u(t_2)) \geq c_1 \|u(t_1) - u(t_2)\|. \quad (7)$$

Hence the curve  $u$  has finite length, and so it is bounded. In view of [8, Theomrem 4.5], there exists the limit  $a := \lim_{s \rightarrow \infty} f(u(s))$ . In addition, we have  $\mathbf{m}_f(a) = 0$ . Let

$$t := \inf\{s : \|u(s)\| > R\}.$$

There are two cases to be considered.

*Case 1:*  $t = \infty$ ; i.e.,  $\|u(s)\| > R$  for all  $s \geq 0$

Since  $\mathbf{m}_f(a) = 0$ , it follows from the inequality (4) that  $f_+(a) = 0$ . Therefore, by (7), we obtain

$$f_+(x) = f_+(x) - f_+(a) \geq c_1 \|u(0) - a\| = c_1 \|x - a\| \geq c_1 d(x, S).$$

*Case 2:*  $t < \infty$

We have  $\|u(t)\| = R$ . Then it follows from (4), (5) and (7) that

$$\begin{aligned} d(x, S) &\leq d(x, u(t)) + d(u(t), S) \\ &\leq \frac{f_+(x) - f_+(u(t))}{c_1} + \frac{(f_+(u(t)))^\alpha}{c_2} \\ &\leq \frac{f_+(x)}{c_1} + \frac{(f_+(x))^\alpha}{c_2}. \end{aligned}$$

In summary, in both cases, we have

$$cd(x, S) \leq f_+(x) + (f_+(x))^\alpha,$$

where  $c := \min\{c_1, c_2\} > 0$ . This, together with (5), implies the required result.

We deduce immediately the following corollary from the proof of Theorem 4 (see also [53, Theorem 3.1]).

**Corollary 1** *Let  $f$  be a continuous semi-algebraic function. Suppose that  $S := \{x \in \mathbb{R}^n : f(x) \leq 0\} \neq \emptyset$  and there exists a constant  $c > 0$  such that*

$$\mathbf{m}_f(x) \geq c, \quad \text{for all } x \in f^{-1}((0, +\infty)).$$

*Then the following linear global error bound holds*

$$cd(x, S) \leq [f(x)]_+, \quad \text{for all } x \in \mathbb{R}^n.$$

Let  $A \in \mathbb{R}^{n \times n}$  be a nonzero symmetric matrix. It is well-known that  $A$  has  $n$  real eigenvalues  $\lambda_i, i = 1, \dots, n$ . Then we make use of the following notation

$$\lambda(A) := \min\{|\lambda_i| : \lambda_i \neq 0, i = 1, \dots, n\}.$$

The following result gives a Hölder-type global error bound result for the zero set of a single quadratic function.

**Corollary 2** Let  $f(x) := \frac{1}{2}x^T Ax + x^T b + c$  be a quadratic function in  $\mathbb{R}^n$ , where  $A \in \mathbb{R}^{n \times n}$  is a nonzero symmetric matrix,  $b$  is a vector in  $\mathbb{R}^n$  and  $c$  is a real number. Let  $\bar{x} \in \mathbb{R}^n$  be such that  $\nabla f(\bar{x}) = A\bar{x} + b = 0$ . Then we have for all  $x \in \mathbb{R}^n$ ,

$$\frac{\sqrt{2\lambda(A)}}{2} d(x, f^{-1}(f(\bar{x}))) \leq |f(x) - f(\bar{x})|^{\frac{1}{2}}.$$

*Proof* We first show the following *gradient inequality* (see [20, Property 6]):

$$\sqrt{2\lambda(A)} |f(x) - f(\bar{x})|^{\frac{1}{2}} \leq \|\nabla f(x)\|, \quad \text{for all } x \in \mathbb{R}^n.$$

Indeed, by the assumption, we have  $b = -A\bar{x}$ . This implies that  $\nabla f(x) = Ax + b = A(x - \bar{x})$  and

$$\begin{aligned} f(x) - f(\bar{x}) &= \frac{1}{2}(x^T Ax - \bar{x}^T A\bar{x}) + (x - \bar{x})^T b \\ &= \frac{1}{2}(x^T Ax - \bar{x}^T A\bar{x}) - (x - \bar{x})^T A\bar{x} \\ &= \frac{1}{2}(x - \bar{x})^T A(x - \bar{x}). \end{aligned}$$

On the other hand, since the matrix  $A$  is symmetric,  $A$  has  $n$  real eigenvalues  $\lambda_i, i = 1, \dots, n$ . Then we can write

$$A = R \operatorname{diag}(\lambda_1, \dots, \lambda_\kappa, 0, \dots, 0) R^T,$$

where  $R \in \mathbb{R}^{n \times n}$  is a suitable orthonormal matrix, and  $\lambda_i, i = 1, \dots, \kappa$ , is the nonzero eigenvalues of  $A$ .

Now let  $x \in \mathbb{R}^n$  and set  $z := R^T(x - \bar{x}) \in \mathbb{R}^n$ . We have

$$\|\nabla f(x)\|^2 = \|A(x - \bar{x})\|^2 = \sum_{i=1}^{\kappa} \lambda_i^2 z_i^2$$

and

$$|f(x) - f(\bar{x})| = \frac{1}{2} |(x - \bar{x})^T A(x - \bar{x})| = \frac{1}{2} \left| \sum_{i=1}^{\kappa} \lambda_i z_i^2 \right| \leq \frac{1}{2} \sum_{i=1}^{\kappa} |\lambda_i| z_i^2.$$

which implies the gradient inequality.

We now consider the continuous semi-algebraic function  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  defined by  $g(x) := |f(x) - f(\bar{x})|^{\frac{1}{2}}$ . Then it easily follows from the above gradient inequality that  $\mathbf{m}_g(x) \geq \frac{\sqrt{2\lambda(A)}}{2}$  for any  $x$  with  $g(x) > 0$ . This, together with Corollary 1, implies the desired result.

*Remark 5* Recall that the authors of [41] (see also [15]) established the following Hölder-type global error bound result for the zero set of a single quadratic function  $f$ : There exists some constant  $c > 0$  such that

$$cd(x, f^{-1}(0)) \leq |f(x)|^{\frac{1}{2}} + |f(x)|, \quad \text{for all } x \in \mathbb{R}^n.$$

However, neither result gives any clue for computing the constant  $c$  in general.

## 6 Proof of the main result

The following lemmas are crucially used in the proof of Theorem 1.

**Lemma 6** *Let  $F = (f_1, \dots, f_p): \mathbb{R}^n \rightarrow \mathbb{R}^p$  be a map of class  $C^1$  and let  $f(x) := \max_{i=1, \dots, p} f_i(x)$ . Then  $f$  is a continuous function and*

$$\mathfrak{m}_f(x) = \min_{\lambda_i \geq 0, \sum_{i \in I} \lambda_i = 1} \left\| \sum_{i \in I} \lambda_i \nabla f_i(x) \right\|,$$

where  $I = I(x) := \{i : f_i(x) = f(x)\}$ .

*Proof* The statement is a consequence of [45, Theorem 3.46(ii)] (cf. also [52, Exercise 8.31]).

**Lemma 7** *Under the assumptions of Theorem 1, there exist some constants  $c > 0$  and  $R > 0$  such that*

$$\mathfrak{m}_f(x) \geq c \quad \text{for all} \quad \|x\| \geq R.$$

*In particular, the function  $f$  is good at infinity.*

*Proof* Suppose that by contradiction there exists a sequence  $\{x^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$  such that

$$\lim_{k \rightarrow \infty} \|x^k\| = \infty \quad \text{and} \quad \lim_{k \rightarrow \infty} \mathfrak{m}_f(x^k) = 0.$$

By definition, there exists a sequence  $\lambda^k := (\lambda_1^k, \dots, \lambda_p^k)$  with  $\lambda_i^k \geq 0$ ,  $\sum_{i \in I(x^k)} \lambda_i^k = 1$  such that

$$\mathfrak{m}_f(x^k) = \left\| \sum_{i \in I(x^k)} \lambda_i^k \nabla f_i(x^k) \right\|.$$

Since the number of subsets of  $\{1, \dots, p\}$  is finite, by taking subsequences if necessary, we may assume that the set  $I(x^k)$  is stable, i.e., there exists  $\tilde{I} \subseteq \{1, \dots, p\}$  such that  $\tilde{I} = I(x^k)$  for all  $k$ . We remark that the function  $\mathfrak{m}_f(x)$  is semi-algebraic. By Lemma 6 and by applying Curve Selection Lemma at infinity (Lemma 1) with the following setup: the set

$$\begin{aligned} A := \{ (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^p : & \lambda_i \geq 0, \sum_{i \in \tilde{I}} \lambda_i = 1, \\ & f_i(x) = f(x) \text{ for } i \in \tilde{I}, \\ & f_i(x) < f(x) \text{ for } i \notin \tilde{I}, \\ & \mathfrak{m}_f(x) = \left\| \sum_{i \in \tilde{I}} \lambda_i \nabla f_i(x) \right\| \} \end{aligned}$$

which is a semi-algebraic set, the sequence  $(x^k, \lambda^k) \in A$  which tends to infinity as  $k \rightarrow \infty$ , and the semi-algebraic function  $x \mapsto \mathfrak{m}_f(x)$ , it follows that there exist a smooth semi-algebraic curve  $\varphi(t) := (\varphi_1(t), \dots, \varphi_n(t))$  and some smooth semi-algebraic functions  $\lambda_i(t), i \in \tilde{I}$ , for  $0 < t \ll 1$ , such that

- (a)  $\lim_{t \rightarrow 0} \|\varphi(t)\| = \infty$ ;
- (b)  $f_i(\varphi(t)) = f(\varphi(t))$  for  $i \in \tilde{I}$ , and  $f_i(\varphi(t)) < f(\varphi(t))$  for  $i \notin \tilde{I}$ ;
- (c)  $\lambda_i(t) \geq 0$  for all  $i \in \tilde{I}$ , and  $\sum_{i \in \tilde{I}} \lambda_i(t) = 1$ .
- (d)  $\mathfrak{m}_f(\varphi(t)) = \left\| \sum_{i \in \tilde{I}} \lambda_i(t) \nabla f_i(\varphi(t)) \right\| \rightarrow 0$  as  $t \rightarrow 0$ .

Let  $J := \{j : \varphi_j \neq 0\}$ . By Condition (a),  $J \neq \emptyset$ . In view of Growth Dichotomy Lemma (Lemma 2), for  $j \in J$ , we can expand the coordinate  $\varphi_j$  in terms of the parameter: say

$$\varphi_j(t) = x_j^0 t^{q_j} + \text{higher order terms in } t,$$

where  $x_j^0 \neq 0$ . From Condition (a), we get  $q_{j_*} := \min_{j \in J} q_j < 0$  for some  $j_* \in J$ . Note that  $\|\varphi(t)\| = ct^{q_{j_*}} + o(t^{q_{j_*}})$  as  $t \rightarrow 0$ , for some  $c > 0$ .

Since  $f_i$  is convenient,  $\Gamma(f_i) \cap \mathbb{R}^J \neq \emptyset$ . Let  $d_i$  be the minimal value of the linear function  $\sum_{j \in J} q_j \kappa_j$  on  $\Gamma(f_i) \cap \mathbb{R}^J$ , and let  $\Delta_i$  be the (unique) maximal face of  $\Gamma(f_i) \cap \mathbb{R}^J$  where the linear function takes this value. Since  $f_i$  is convenient,  $d_i < 0$  and  $\Delta_i$  is a face of  $\Gamma_\infty(f_i)$ . Note that  $f_{i, \Delta_i}$  does not depend on  $x_j$  for all  $j \notin J$ . By a direct calculation, then

$$f_i(\varphi(t)) = f_{i, \Delta_i}(x^0) t^{d_i} + \text{higher order terms in } t,$$

where  $x^0 := (x_1^0, \dots, x_n^0)$  with  $x_j^0 = 1$  for  $j \notin J$ .

Let  $I := \{i \in \tilde{I} : \lambda_i \neq 0\}$ . It follows from Condition (c) that  $I \neq \emptyset$ . For  $i \in I$ , expand the coordinate  $\lambda_i$  in terms of the parameter: say

$$\lambda_i(t) = \lambda_i^0 t^{\theta_i} + \text{higher order terms in } t,$$

where  $\lambda_i^0 \neq 0$ .

For  $i \in I$  and  $j \in J$  we have

$$\frac{\partial f_i}{\partial x_j}(\varphi(t)) = \frac{\partial f_{i, \Delta_i}}{\partial x_j}(x^0) t^{d_i - q_j} + \text{higher order terms in } t.$$

It implies that

$$\begin{aligned} \sum_{i \in I} \lambda_i(t) \frac{\partial f_i}{\partial x_j}(\varphi(t)) &= \sum_{i \in I} \left( \lambda_i^0 \frac{\partial f_{i, \Delta_i}}{\partial x_j}(x^0) t^{d_i + \theta_i - q_j} + \text{higher order terms in } t \right) \\ &= \left( \sum_{i \in I'} \lambda_i^0 \frac{\partial f_{i, \Delta_i}}{\partial x_j}(x^0) \right) t^{\ell - q_j} + \text{higher order terms in } t, \end{aligned}$$

where  $\ell := \min_{i \in I} (d_i + \theta_i)$  and  $I' := \{i \in I : d_i + \theta_i = \ell\} \neq \emptyset$ .

There are two cases to be considered.

*Case 1:*  $\ell \leq q_{j_*} := \min_{j \in J} q_j$

We deduce from Condition (d) that

$$\sum_{i \in I'} \lambda_i^0 \frac{\partial f_{i, \Delta_i}}{\partial x_j}(x^0) = 0, \quad \text{for all } j \in J,$$

which yields

$$\sum_{i \in I'} \lambda_i^0 \frac{\partial f_{i, \Delta_i}}{\partial x_j}(x^0) = 0, \quad \text{for all } j = 1, 2, \dots, n,$$

because  $f_{i, \Delta_i}$  does not depend on  $x_j$  for all  $j \notin J$ . It implies easily that

$$\text{rank} \left( x_j^0 \frac{\partial f_{i, \Delta_i}}{\partial x_j}(x^0) \right)_{i \in I', 1 \leq j \leq n} < \#I'.$$

Since the map  $F = (f_1, \dots, f_p)$  is non-degenerate at infinity, there exists an index  $i_0 \in I'$  such that  $f_{i_0, \Delta_{i_0}}(x^0) \neq 0$ . Then, by Condition (b), we have for all  $i \in \tilde{I}$ ,

$$f(\varphi(t)) = f_i(\varphi(t)) = f_{i_0}(\varphi(t)) = f_{i_0, \Delta_{i_0}}(x^0)t^{d_{i_0}} + \text{higher order terms in } t.$$

By taking the derivative in  $t$  of the function  $(f \circ \varphi)(t)$ , we deduce that

$$\frac{d(f \circ \varphi)(t)}{dt} = \frac{d(f_i \circ \varphi)(t)}{dt} = \left\langle \nabla f_i(\varphi(t)), \frac{d\varphi(t)}{dt} \right\rangle, \quad \text{for all } i \in \tilde{I}.$$

By Condition (c), then

$$\frac{d(f \circ \varphi)(t)}{dt} = \sum_{i \in \tilde{I}} \lambda_i(t) \frac{d(f_i \circ \varphi)(t)}{dt} = \left\langle \sum_{i \in \tilde{I}} \lambda_i(t) \nabla f_i(\varphi(t)), \frac{d\varphi(t)}{dt} \right\rangle.$$

Thus

$$\left| \frac{d(f \circ \varphi)(t)}{dt} \right| \leq \left\| \sum_{i \in \tilde{I}} \lambda_i(t) \nabla f_i(\varphi(t)) \right\| \left\| \frac{d\varphi(t)}{dt} \right\| = \mathbf{m}_f(\varphi(t)) \left\| \frac{d\varphi(t)}{dt} \right\|,$$

which implies that

$$\mathbf{m}_f(\varphi(t)) \geq c' t^{d_{i_0} - q_{j^*}} + \text{higher order terms in } t,$$

for some  $c' > 0$ . But this inequality contradicts Condition (d) since we know that

$$d_{i_0} \leq d_{i_0} + \theta_{i_0} = \ell \leq q_{j^*}.$$

*Case 2:*  $\ell > q_{j^*} := \min_{j \in J} q_j$

It follows from Condition (c) that  $\theta_i \geq 0$  for all  $i \in I$  and  $\theta_i = 0$  for some  $i \in I$ . Without loss of generality, we may assume that  $1 \in I$  and  $\theta_1 = 0$ . Since  $f_1$  is convenient, for any  $j = 1, \dots, n$ , there exists a natural number  $m_j \geq 1$  such that  $m_j e_j \in \Gamma_\infty(f_1)$ . Then it is clear that

$$q_j m_j \geq d_1, \quad \text{for all } j \in J.$$

On the other hand, we have

$$d_1 = d_1 + \theta_1 \geq \min_{i \in I} (d_i + \theta_i) = \ell.$$

Therefore

$$q_{j^*} m_{j^*} \geq d_1 \geq \ell > q_{j^*}.$$

Since  $q_{j^*} = \min_{j \in J} q_j < 0$ , it implies that  $m_{j^*} < 1$ , which is a contradiction.

**Corollary 3** *Under the assumptions of Theorem 1, there exist some positive constants  $c, \delta$  and  $\alpha$  such that the following Hölder-type error bound “near to  $S := \{x \in \mathbb{R}^n : \mathbf{m}_f(x) = 0\}$ ” holds*

$$cd(x, S) \leq [\mathbf{m}_f(x)]^\alpha \quad \text{for all } x \in \mathbf{m}_f \leq \delta.$$

*Proof* By Lemma 7, there exist some constants  $c_1 > 0$  and  $R > 0$  such that

$$\mathbf{m}_f(x) \geq c_1 \quad \text{for all } \|x\| \geq R.$$

Hence there is no sequence  $x^k \rightarrow \infty$  such that  $\mathbf{m}_f(x) \rightarrow 0$ . By Lemma 4, there exist some constants  $c > 0, \delta > 0$ , and  $\alpha > 0$  such that

$$cd(x, S) \leq [\mathbf{m}_f(x)]^\alpha \quad \text{for all } x \in \mathbf{m}_f \leq \delta.$$

*Remark 6* Lemma 7 was proved by another method by Hà [24] (see also [11, Proposition 3.4]) for a single polynomial function; i.e., for the case where  $p = 1$ .

Before proving Theorem 1 which establishes that a Hölder-type global error bound holds with an explicit exponent, we recall an error bound result on a bounded region.

**Lemma 8** *Let  $S$  denote the set of  $x$  in  $\mathbb{R}^n$  satisfying  $f_1(x) \leq 0, \dots, f_p(x) \leq 0$ , where each  $f_i$  is a real polynomial. Let  $R$  be a positive number such that  $S$  contains an element  $x$  with  $\|x\| \leq R$ . Then, there exists a constant  $c > 0$  such that*

$$cd(x, S) \leq [f(x)]_+^{\frac{2}{\mathcal{H}(2d, n, p)}} \quad \text{for all } x \text{ with } \|x\| \leq R. \quad (8)$$

Here  $f(x) := \max_{i=1, \dots, p} f_i(x)$  and  $d := \max_{i=1, \dots, p} \deg f_i$ .

*Proof* See [36, Corollary 3.8].

Now, we are in position to finish the proof of Theorem 1.

*Proof (Proof of Theorem 1)* By Lemma 7, the continuous semi-algebraic function  $f: \mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto f(x) := \max_{i=1, \dots, p} f_i(x)$ , is good at infinity. Then the proof follows on the same lines as that of Theorem 4, by using the inequality (8) in Lemma 8 instead of the inequality (5). We omit the details.

*Remark 7* Theorem 1 was obtained by Hà [24] for a single polynomial function; however, there was no explicit formula given for computing the exponents  $\alpha$  and  $\beta$ . Again, the present proof is different from that in the above cited paper.

## 7 Examples

In this section, we give some examples which illustrate Theorem 1.

Denote the convex hull of a set of points  $0, a_1, \dots, a_m \in \mathbb{R}^n$  by  $\Gamma\{a_1, \dots, a_m\}$ .

*Example 1* Consider the following polynomial map

$$F = (f_1, f_2): \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (x + y, x^2 + y^2 - 1).$$

Note that  $f_1, f_2$  are convenient and  $S$  is the half-disk  $\{x + y \leq 0, x^2 + y^2 \leq 1\}$ . The Newton polyhedra at infinity of  $f_1$  and  $f_2$ , are the triangles  $\Gamma(f_1) = \Gamma\{(1, 0), (0, 1)\}$  and  $\Gamma(f_2) = \Gamma\{(2, 0), (0, 2), (0, 0)\}$ , respectively. The Minkowski sum

$$\Gamma(F) = \Gamma(f_1) + \Gamma(f_2) = \Gamma\{(3, 0), (0, 3), (0, 0)\}$$

is again a triangle. Then  $\Gamma_\infty(F)$  has three faces which are  $\Delta^1 := \{(3, 0)\} = \{(1, 0)\} + \{(2, 0)\}$ ,  $\Delta^2 := \{(0, 3)\} = \{(0, 1)\} + \{(0, 2)\}$ , and  $\Delta^3 := \Gamma\{(3, 0), (0, 3)\} = \Gamma\{(1, 0), (0, 1)\} + \Gamma\{(2, 0), (0, 2)\}$ . So we have  $f_{\Delta^1} = (x, x^2)$ ,  $f_{\Delta^2} = (y, y^2)$ , and  $f_{\Delta^3} = (x + y, x^2 + y^2)$ . It is clear that the following corresponding matrices

$$M_{\Delta^1} := \begin{pmatrix} x & 0 & x & 0 \\ 2x^2 & 0 & 0 & x^2 \end{pmatrix}, \quad M_{\Delta^2} := \begin{pmatrix} 0 & y & y & 0 \\ 0 & 2y^2 & 0 & y^2 \end{pmatrix}, \quad M_{\Delta^3} := \begin{pmatrix} x & y & x + y & 0 \\ 2x^2 & 2y^2 & 0 & x^2 + y^2 \end{pmatrix}$$

have rank 2 on  $(\mathbb{R} \setminus \{0\})^2$ . Hence  $F$  is non-degenerate at infinity. The reader may easily check the following Hölder-type global error bound

$$cd(x, S) \leq [f(x)]_+^{\frac{1}{2}} + [f(x)]_+ \quad \text{for all } x \in \mathbb{R}^n,$$



and for some  $c > 0$ . The exponent  $\alpha = \frac{1}{2}$  here can be also obtained from [41, Corollary 16.14] by restricting on the case of systems of one linear and one convex quadratic inequality.

In general, it is not easy to verify directly whenever a system of polynomials has Hölder-type global error bounds or not. However, it can be done by checking the condition of non-degeneracy at infinity.

*Example 2* Let

$$F = (f_1, f_2): \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad (x, y, z) \mapsto (x^2 + y^2 + z^2, x + y + z^3).$$

We have

$$\begin{aligned} \Gamma(f_1) &= \Gamma\{(2, 0, 0), (0, 2, 0), (0, 0, 2)\}, \\ \Gamma(f_2) &= \Gamma\{(1, 0, 0), (0, 1, 0), (0, 0, 3)\}. \end{aligned}$$

Then  $f_1, f_2$  are convenient and

$$\Gamma(F) = \Gamma\{(3, 0, 0), (0, 3, 0), (0, 0, 5), (2, 0, 3), (0, 2, 3)\}.$$

Hence

$$\Gamma_\infty(F) = \{\Delta^1, \dots, \Delta^{13}\},$$

with

- (i)  $\Delta^1 = \Gamma\{(3, 0, 0), (0, 3, 0), (2, 0, 3), (0, 2, 3)\} = \Gamma\{(2, 0, 0), (0, 2, 0)\} + \Gamma\{(1, 0, 0), (0, 1, 0), (0, 0, 3)\},$
- (ii)  $\Delta^2 = \Gamma\{(2, 0, 3), (0, 2, 3), (0, 0, 5)\} = \Gamma\{(2, 0, 0), (0, 2, 0), (0, 0, 2)\} + \{(0, 0, 3)\},$
- (iii)  $\Delta^3 = \Gamma\{(3, 0, 0), (0, 3, 0)\} = \Gamma\{(2, 0, 0), (0, 2, 0)\} + \Gamma\{(1, 0, 0), (0, 1, 0)\},$
- (iv)  $\Delta^4 = \Gamma\{(2, 0, 3), (0, 2, 3)\} = \Gamma\{(2, 0, 0), (0, 2, 0)\} + \{(0, 0, 3)\},$
- (v)  $\Delta^5 = \Gamma\{(2, 0, 3), (0, 0, 5)\} = \Gamma\{(2, 0, 0), (0, 0, 2)\} + \{(0, 0, 3)\},$
- (vi)  $\Delta^6 = \Gamma\{(3, 0, 0), (2, 0, 3)\} = \{(2, 0, 0)\} + \Gamma\{(1, 0, 0), (0, 0, 3)\},$
- (vii)  $\Delta^7 = \Gamma\{(0, 2, 3), (0, 0, 5)\} = \Gamma\{(0, 2, 0), (0, 0, 2)\} + \{(0, 0, 3)\},$
- (viii)  $\Delta^8 = \Gamma\{(0, 3, 0), (0, 2, 3)\} = \{(0, 2, 0)\} + \Gamma\{(0, 1, 0), (0, 0, 3)\},$
- (ix)  $\Delta^9 = \{(3, 0, 0)\} = \{(2, 0, 0)\} + \{(1, 0, 0)\},$
- (x)  $\Delta^{10} = \{(0, 3, 0)\} = \{(0, 2, 0)\} + \{(0, 1, 0)\},$
- (xi)  $\Delta^{11} = \{(2, 0, 3)\} = \{(2, 0, 0)\} + \{(0, 0, 3)\},$
- (xii)  $\Delta^{12} = \{(0, 2, 3)\} = \{(0, 2, 0)\} + \{(0, 0, 3)\},$
- (xiii)  $\Delta^{13} = \{(0, 0, 5)\} = \{(0, 0, 2)\} + \{(0, 0, 3)\},$

By computation, it is not hard to show that the corresponding matrices  $M_{\Delta^j}$  have rank 2 on  $(\mathbb{R} \setminus \{0\})^3$  for  $j = 1, \dots, 13$ . Hence  $F$  is non-degenerate at infinity. By Theorem 1,  $F$  has a Hölder-type global error bound with the exponents  $\alpha = \frac{2}{2d(12d-3)^{n+p-1}} = \frac{1}{3 \times 33^4}$  and  $\beta = 1$ .

*Example 3* Let

$$F = (f_1, f_2): \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (x^2 - y^2, x - y).$$

It is clear that

$$\begin{aligned} \Gamma(f_1) &= \Gamma\{(2, 0), (0, 2)\}, \\ \Gamma(f_2) &= \Gamma\{(1, 0), (0, 1)\}, \\ \Gamma(F) &= \Gamma(f_1) + \Gamma(f_2) = \Gamma\{(3, 0), (0, 3)\}. \end{aligned}$$

Consider the edge  $\Delta := \Gamma\{(3, 0), (0, 3)\} = \Gamma\{(2, 0), (0, 2)\} + \Gamma\{(1, 0), (0, 1)\}$  of  $\Gamma_\infty(F)$ , then  $F_\Delta = F$  and

$$M_\Delta = \begin{pmatrix} 2x^2 - 2y^2 & x^2 - y^2 & 0 \\ x & -y & 0 & x - y \end{pmatrix}.$$

It is clear that  $\text{rank} M_\Delta = 1$  when  $x = y \neq 0$ , so the condition of non-degeneracy at infinity is not satisfied. However, by a small perturbation  $F_\epsilon = F + (0, \epsilon x) = (x^2 - y^2, x - y + \epsilon x)$ , for any non-zero small  $\epsilon$ , we still have  $\Gamma_\infty(F_\epsilon) = \Gamma_\infty(F)$  and the reader may check that the new polynomial map  $F_\epsilon$  is non-degenerate at infinity. The exponents of the Hölder-type global error bound of  $F_\epsilon$  will be  $\alpha = \frac{2}{2d(12d-3)^{n+p-1}} = \frac{1}{2 \times 21^3}$  and  $\beta = 1$ .

## 8 Applications

In this section, we describe various applications of the error bound results obtained in the previous sections.

### 8.1 0 – 1 integer feasibility problem

Let  $F: \mathbb{R}^n \rightarrow \mathbb{R}^p$  and  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  be polynomial maps. Consider the following 0 – 1 integer feasibility problem (see [40]):

$$F(x) \leq 0, \quad g(x) = 0, \quad x_i = 0 \text{ or } 1, \quad i = 1, \dots, n.$$

Equivalently, we may consider the following system:

$$F(x) \leq 0, \quad g(x) = 0, \quad x_i(x_i - 1) = 0, \quad i = 1, \dots, n.$$

Assume that the solution set  $S$  of the problem is not empty, and set

$$r(x) := \|[f(x)]_+\| + |g(x)| + \sum_{i=1}^n |x_i(x_i - 1)|.$$

It is clear that  $r(x)$  is a nonnegative, continuous semi-algebraic function. Moreover,  $r(x)$  is proper, i.e.,  $r(x^k) \rightarrow \infty$  for any sequence  $x^k \rightarrow \infty$ . Hence  $r(x)$  satisfies the Palais-Smale condition at any level  $t \geq 0$ . By Theorem 3, there exist some constants  $c > 0, \alpha > 0$ , and  $\beta > 0$  such that for all  $x \in \mathbb{R}^n$ , we have

$$cd(x, S) \leq [r(x)]^\alpha + [r(x)]^\beta.$$

In fact, this Hölder-type global error bound still holds if  $F$  and  $g$  are analytic mapping so this generalizes Theorem 5.6 in [40], which gives a Hölder-type error bound in the compact setting.

### 8.2 Partition problem

The partition problem asks whether an integer sequence  $a_1, \dots, a_n$  can be partitioned, i.e., whether there exists  $x \in \{\pm 1\}^n$  such that  $\sum_{j=1}^n a_j x_j = 0$ . This problem is known to be NP-complete (see [22]). If the infimum  $f^*$  of the polynomial  $f := (\sum_{j=1}^n a_j x_j)^2 + \sum_{j=1}^n (x_j^2 - 1)^2$  on  $\mathbb{R}^n$  is equal to 0, a global minimizer is  $\pm 1$ -valued and thus provides a partition of the sequence.

We leave the reader verifying that  $f$  is convenience and non-degenerate at infinity. It follows from Theorem 1.1 in [16] that  $f$  attains its infimum on  $\mathbb{R}^n$ . We have  $d = 4$  and  $p = 1$ , so  $\mathcal{H}(2d, n, p) = \mathcal{H}(8, n, 1) = 8(45)^n$ . Assume that  $f^* = 0$ ; then  $S = \{x \in \mathbb{R}^n : f(x) \leq 0\} = \{x \in \mathbb{R}^n : f(x) = 0\} \neq \emptyset$ . By Theorem 1, the following Hölder-type global error bound holds

$$cd(x, S) \leq [f(x)]^{\frac{1}{8(45)^n}} + f(x) \quad \text{for all } x \in \mathbb{R}^n.$$

### 8.3 Growth rate of the objective function in a polynomial optimization program

Let  $f_0$  and  $f_1, \dots, f_p : \mathbb{R}^n \rightarrow \mathbb{R}$  be polynomial functions in  $n$  real variables. Set

$$S := \{x \in \mathbb{R}^n : f_1(x) \leq 0, \dots, f_p(x) \leq 0\}.$$

Assume that  $S$  is non empty. Let us consider the following constrained optimization problem

$$f^* := \inf_{x \in S} f_0(x) \quad \text{such that } x \in S$$

of minimizing  $f_0$  over  $S$ . Under the condition of convenience and non-degeneracy at infinity, we have proved in [16] that if  $f_0$  is bounded from below on  $S$ , then  $f_0$  attains its infimum on  $S$ . Hence Theorem 1 can be applied to the solution set of this nonlinear polynomial program to obtain a growth property of the objective function.

**Corollary 4** *Assume that  $f_0$  is bounded from below on  $S$  and that the map  $(f_0, f_1, \dots, f_p)$  is convenient and non-degenerate at infinity, then  $f_0$  attains its infimum on  $S$ . Let*

$$A := \{x \in S : f_0(x) = f^*\} \neq \emptyset$$

*be the set of globally optimal solutions of  $f$  on  $S$ . Then there exists a constant  $c > 0$  such that*

$$cd(x, A) \leq [f_0(x) - f^*]^{\frac{2}{\mathcal{H}(2d, n, p)}} + [f_0(x) - f^*] \quad \text{for all } x \in S.$$

### 8.4 Global Hölderian stability for set-valued maps

Let  $F = (f_1, \dots, f_p) : \mathbb{R}^n \rightarrow \mathbb{R}^p$  be a polynomial map. We define the set-valued map  $S : \mathbb{R}^p \rightrightarrows \mathbb{R}^n$  by

$$S(y) := \{x \in \mathbb{R}^n : f_i(x) - y_i \leq 0, i = 1, \dots, p\} \quad \text{for } y := (y_1, \dots, y_p).$$

Then we have the following global Hölderian property of the set-valued map  $S$ .

**Corollary 5** *Assume that  $F$  is convenient and non-degenerate at infinity. Then there exists a positive constant  $c$  such that*

$$S(y) \subseteq S(0) + c(\|y\|^{\frac{2}{\mathcal{H}(2d, n, p)}} + \|y\|)\mathbb{B} \quad \text{for all } y \in \mathbb{R}^p,$$

*where  $\mathbb{B}$  denotes the closed unit Euclidean ball centered at the origin in  $\mathbb{R}^n$ .*

*Proof* Let us define the function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  by  $f(x) := \max_{i=1, \dots, p} f_i(x)$ . In view of Theorem 1, there exists a constant  $c' > 0$  such that

$$c'd(x, S(0)) \leq [f(x)]_+^{\frac{2}{\mathcal{H}(2d, n, p)}} + [f(x)]_+ \quad \text{for all } x \in \mathbb{R}^n.$$

Let  $y := (y_1, \dots, y_p)$  be arbitrary in  $\mathbb{R}^p$ . It suffices to show that

$$c'd(x, S(0)) \leq \|y\|^{\frac{2}{\mathcal{H}(2d, n, p)}} + \|y\| \quad \text{for all } x \in S(y).$$

In fact, take any  $x \in S(y)$ . Then

$$[f(x)]_+ = \max\{f(x), 0\} \leq \max\{y_1, \dots, y_p, 0\} \leq \|y\|.$$

Therefore

$$c'd(x, S(0)) \leq [f(x)]_+^{\frac{2}{\mathcal{H}(2d, n, p)}} + [f(x)]_+ \leq \|y\|^{\frac{2}{\mathcal{H}(2d, n, p)}} + \|y\|,$$

which completes the proof.

**Acknowledgements** This research was performed while the authors had been visiting at Vietnam Institute for Advanced Study in Mathematics (VIASM). The authors would like to thank the Institute for hospitality and support.

<sup>†</sup>These authors were partially supported by Vietnam National Foundation for Science and Technology Development (NAFOS-TED) grant 101.04-2014.23 and the Vietnam Academy of Science and Technology (VAST)

<sup>‡</sup>This author was partially supported by Vietnam National Foundation for Science and Technology Development (NAFOS-TED) grant 101.04-2013.07

## References

1. Arnold, V.I., Gusein-Zade, S., Varchenko, A.N.: Singularities of differentiable maps. Monogr. Math., Vol. I and II, Birkhäuser, Basel, (1985).
2. Auslender, A.A., Crouzeix, J.-P.: Global regularity theorem. Math. Oper. Res., 13, 243-253, (1988).
3. Auslender, A.A., and Crouzeix, J.-P.: Well behaved asymptotical convex functions. Ann. Inst. H. Poincaré, Anal. Non Linéaire 6, 101-121 (1989)
4. Auslender, A.A., Cominetti, R., Crouzeix, J.-P.: Convex functions with unbounded level sets. SIAM J. Optim. 3, 669-687 (1993).
5. Benedetti, R., Risler, J.: Real algebraic and semi-algebraic sets. Hermann, (1991).
6. Bierstone, E., Milman, P.: Semianalytic and subanalytic sets. Inst. Hautes Études Sci. Publ. Math. 67 , 5-42, (1988).
7. Bochnak, J., Coste, M., Roy, M.-F.: Real algebraic geometry. Vol. 36, Springer, (1998).
8. Bolte, J., Daniilidis, A., Lewis, A.S.: The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. SIAM J. Optim. 17 , no. 4, 1205-1223, (2007).
9. Borwein, J. M., Li, G., Yao, L.: Analysis of the convergence rate for the cyclic projection algorithm applied to basic semi-algebraic convex sets. SIAM J. Optim. 24, 498-527, (2014).
10. Borwein, J. M., Preiss, D.: A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions. Trans. Amer. Math. Soc. 303, 517-527, (1987).
11. Broughton, S.: Milnor number and topology of polynomial hypersurfaces. Invent. Math., 92, 217-241, (1988).
12. Clarke, F.H.: Optimization and nonsmooth analysis. New York et al., John Wiley & Sons (1983).
13. Clarke, F.H., Ledyaev, Yu.S., Stern, R.J., Wolenski, P.R.: Nonsmooth analysis and control theory. Springer-Verlag New York, Inc, (1998).
14. Corvellec, J.-N., Motreanu, V.V.: Nonlinear error bounds for lower semicontinuous functions on metric spaces. Math. Program. 114 , no. 2, Ser. A, 291-319, (2008).
15. Dinh, S.T., Hà, H.V., Thao, N.T.: Lojasiewicz inequality for polynomial functions on non compact domains. International Journal of Mathematics, Vol. 23, No. 4, 1250033 (28 pages), (2012) DOI: 10.1142/S0129167X12500334.
16. Dinh, S.T., Hà, H.V., Phạm, T.S.: A Frank-Wolfe type theorem for nondegenerate polynomial programs. Mathematical Programming SERIES A, DOI 10.1007/s10107-013-0732-2.
17. Dinh, S.T., Hà, H.V., Phạm, T.S., Thao, N.T.: Global Lojasiewicz-type inequality for non-degenerate polynomial maps. J. Math. Anal. Appl. 410, no. 2, 541-560, (2014).
18. van den Dries, L., Miller, C.: Geometric categories and o-minimal structures. Duke Math. J., 84, 497-540, (1996).
19. Ekeland, I.: Nonconvex minimization problems. Bull. A.M.S., No. 1, 443-474, (1979).
20. Forti, M., Tesi, A.: The Lojasiewicz exponent at an equilibrium point of a standard CNN is 1/2. Internat. J. Bifur. Chaos Appl. Sci. Engrg. 16 , no. 8, 2191-2205, (2006).
21. Gaffney, T.: Integral closure of modules and Whitney equisingularity. Invent. Math., 107, 301-322, (1992).
22. Garey, M.R., Johnson, D.S.: Computers and intractability: A guide to the theory of NP-completeness. San Francisco, W.H. Freeman & Company, Publishers, (1979).
23. Hà, H.V., Duc, N.H.: Lojasiewicz inequality at infinity for polynomials in two real variables. Math. Z., 266, 243-264, (2010).
24. Hà, H.V.: Global Hölderian error bound for non-degenerate polynomials. SIAM J. Optim., 23:2, 917-933 (2013).
25. Hoffman, A.J.: On approximate solutions of linear inequalities. Journal of Research of the National Bureau of Standards, 49, 263-265, (1952).
26. Hörmander, L.: On the division of distributions by polynomials. Ark. Mat. 3 N. 53, 555-568, (1958).
27. Ioffe, A.D.: Metric regularity and subdifferential calculus. Russian Math. Surveys, 55, No.3, 501-558, (2000).
28. Ji, S., Kollár, J., Shiffman, B.: A global Lojasiewicz inequality for algebraic varieties. Trans. Amer. Math. Soc., 329, No. 2, 813-818, (1992).
29. Khovanskii, A.G.: Newton polyhedra and toroidal varieties. Funct. Anal. Appl., 11, 289-296 (1978).
30. Khovanskii, A.G.: Fewnomials. Translated from the Russian by Smilka Zdravkovska. Translations of Mathematical Monographs, 88. American Mathematical Society, Providence, RI, (1991).

31. Klatte, D.: Hoffman's error bound for systems of convex inequalities. *Mathematical Programming with data perturbations*, 185-199, *Lecture Notes in Pure and Appl. Math.*, 195, Dekker New York, (1998).
32. Klatte, D., Li, A.: Asymptotic constraint qualifications and global error bounds for convex inequalities. *Math. Program.*, 84, 137-140 (1999).
33. Kouchnirenko, A.G.: Polyhedres de Newton et nombre de Milnor. *Invent. math.*, 32, 1-31 (1976).
34. Lemaire, B.: Bonne position, conditionnement, et bon comportement asymptotique. *Sém. Anal. Convexe* 22, Exp. No. 5, 12 pp, (1992).
35. Lewis, A.S., Pang, J.S.: Error bounds for convex inequality systems. *Generalized convexity, generalized monotonicity: recent results* (Luminy, 1996), 75-110, *Nonconvex Optim. Appl.*, 27, Kluwer Acad. Publ., Dordrecht, (1998).
36. Li, C., Mordukhovich, B.S., Phạm, T.S.: New fractional error bounds for polynomial systems with applications to Hölderian stability in optimization and spectral theory of tensors. *Math. Program.*, (Accepted).
37. Li, G., Mordukhovich, B.S.: Hölder metric subregularity with applications to proximal point method. *SIAM J. Optim.* 22, 1655-1684, (2012).
38. Li, W.: Error bounds for piecewise convex quadratic programs and applications. *SIAM J. Control Optim.* 33, 1510-1529, (1995).
39. Łojasiewicz, S.: Division d'une distribution par une fonction analytique de variables réelles. *C. R. Acad. Sci. Paris*, **246**, 683-686, (1958).
40. Luo, Z.-Q., Pang, J.S.: Error bounds for analytic systems and their applications. *Math. Program.*, 67, pp. 1-28 (1994).
41. Luo, Z.-Q., Sturm, J.F.: Error bound for quadratic systems, in *High Performance Optimization*. H. Frenk, K. Roos, T. Terlaky, and Zhang, eds., Kluwer, Dordrecht, The Netherlands, 383-404, (2000).
42. Mangasarian, O.L.: A condition number for differentiable convex inequalities. *Math. Oper. Res.*, 10, 175-179 (1985).
43. Miller, C.: Exponentiation is hard to avoid. *Proc. Amer. Math. Soc.* 122, 257-259 (1994).
44. Milnor, J.: Singular points of complex hypersurfaces. *Annals of Mathematics Studies* **61**, Princeton University Press, (1968).
45. Mordukhovich, B.S.: Variational analysis and generalized differentiation, I: Basic Theory, II: Applications. Springer, Berlin, (2006).
46. Némethi, A., Zaharia, A.: Milnor fibration at infinity. *Indag. Math.*, **3**, 323-335 (1992).
47. Ngai, H.V., Thera, M.: Error bounds for differentiable convex inequality systems in Banach spaces. *Math. Program., Ser. B*, 104, No.2-3, 465-482 (2005).
48. Oka, M.: Non-degenerate complete intersection singularity. *Actualités Mathématiques*, Hermann, Paris, (1997).
49. Pang, J.S.: Error bounds in mathematical programming. *Math. Program., Ser.B*, 79, 299-332 (1997).
50. Penot, J.P.: Well-behavior, well-posedness and nonsmooth analysis. *Proceedings of the 4th International Conference on Mathematical Methods in Operations Research and 6th Workshop on Well-posedness and Stability of Optimization Problems* (Sozopol, 1997). *Pliska Stud. Math. Bulgar.* 12, 141-190 (1998).
51. Robinson, S.M.: An application of error bounds for convex programming in a linear space. *SIAM J. Control* 13, 271-273, (1975).
52. Rockafellar, R.T., Wets, R.: Variational analysis. *Grundlehren Math. Wiss.*, **317**, Springer, New York, (1998).
53. Wu, Z., Ye, J.J.: Sufficient conditions for error bounds. *SIAM J. Optim.* 12, 421-435, (2001).